DOE/SC-0066

GENOMES to LIFE

BIOLOGICAL SOLUTIONS FOR ENERGY CHALLENGES

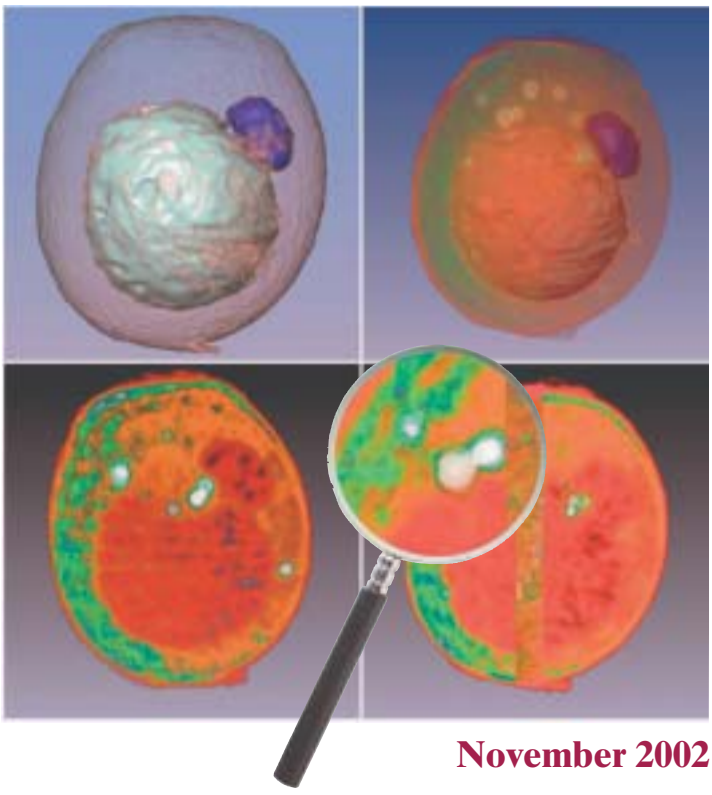
U.S. DEPARTMENT OF ENERGY

INNOVATIVE APPROACHES ALONG UNCONVENTIONAL PATHS

DOEGenomesToLife.org/technology

# Report on the Imaging Workshop for the Genomes to Life Program

## April 16–18, 2002



Sponsored by the
Office of Advanced Scientific Computing Research
and
Office of Biological and Environmental Research
of the
U.S. Department of Energy
Office of Science

November 2002

## Genomes to Life Program

### *Contacts for program information*

**Marvin Frazier**
**U.S. Department of Energy (SC-72)**
**Office of Biological and Environmental Research**
**301/903-5468, Fax: 301/903-8521**
**marvin.frazier@science.doe.gov**

**Gary Johnson**
**U.S. Department of Energy (SC-30)**
**Office of Advanced Scientific Computing Research**
**301/903-5800, Fax: 301/903-7774**
**garyj@er.doe.gov**

### *Publications*

**Documents, meeting reports, and image gallery via the Web:**

- http://DOEGenomesToLife.org

**Copies of this report and other Genomes to Life publications:**

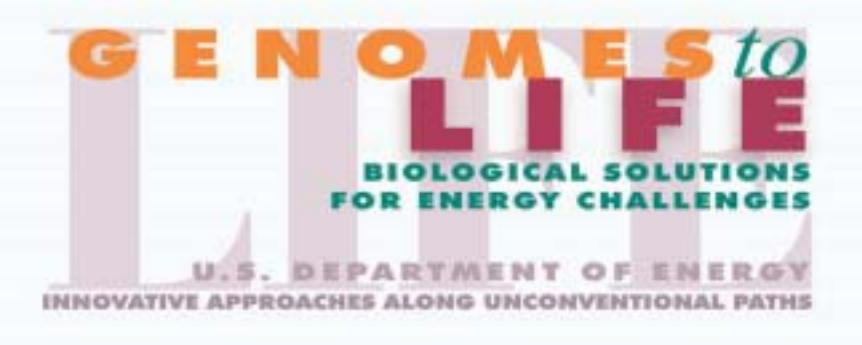
- Human Genome Management Information System
  Oak Ridge National Laboratory
  1060 Commerce Park, MS 6480
  Oak Ridge, TN 37830

  865/576-6669, Fax: 865/574-9888, mansfieldbk@ornl.gov

**Electronic versions of this report and additional information:**

- http://DOEGenomesToLife.org/technology

DOE/SC-0066




*DOEGenomesToLife.org/technology*

# Report on the Imaging Workshop for the Genomes to Life Program

## Charlotte, North Carolina
## April 16–18, 2002

**Co-Chairs**
**Steven Colson, Pacific Northwest National Laboratory**
**Damir Sudar, Lawrence Berkeley National Laboratory**

**Report prepared by the**
**Office of Advanced Scientific Computing Research**
**and**
**Office of Biological and Environmental Research**
**of the**
**U.S. Department of Energy**
**Office of Science**

**November 2002**

# Table of Contents

# List of Images

# Preface

This report is a result of the Imaging Workshop for the Genomes to Life (GTL) program held April 16–19, 2002, in Charlotte, North Carolina. The meeting was sponsored by the Office of Biological and Environmental Research and the Office of Advanced Scientific Computing Research of the U.S. Department of Energy's (DOE) Office of Science.

The purpose of the workshop was to project a broad vision for future needs and determine the value of imaging to GTL program research. The workshop included four technical sessions with plenary lectures on biology and technology perspectives and technical presentations on needs and approaches as they related to the following areas of the GTL program:

1. Molecular machines (protein complexes)
2. Intracellular and cellular structure, function, and processes
3. Multicellular: Monoclonal and heterogeneous multicellular systems, cell-cell signaling, and model systems
4. Cells in situ and in vivo: Bacteria in the natural environment, microenvironment, and in vivo systems

More than 60 attended the workshop, including scientists from most of the DOE national laboratories and partnering universities and GTL program managers. (A list of the steering committee and DOE advisors is included in Appendix A and the workshop agenda in Appendix B.) Participants were divided into seven writing teams that met after the technical sessions to draft this report, which was originally divided into the following sections:

- Executive Summary
- Molecular Engines
- Intracellular and Cellular
- Multicellular
- Cells In Vivo
- Image Data Computational Infrastructure
- Probe Development

The sections were compiled and edited at Pacific Northwest National Laboratory. Oak Ridge National Laboratory did the final editing and prepared the report for publication. Oak Ridge Institute for Science and Education made all arrangements for the workshop.

# Executive Summary

*Thought is impossible without an image.*
**— Aristotle, 325 B.C.E.**

The overall goal of the U.S. Department of Energy's (DOE) Genomes to Life (GTL) program is to understand the composition and function of the biochemical networks and pathways that carry out the essential processes of living microbial organisms. Such understanding is essential for DOE to more effectively address its missions in energy security, carbon management, and environmental cleanup. Imaging of microbial organisms is an essential enabling component of GTL because it provides a method for linking genomic information to function. Imaging aids understanding of how cell function changes with time and environment. Innovations in imaging coupled with computational advances will accelerate scientific discovery and enable biological solutions to energy challenges.

GTL has four main goals:

**Goal 1:** Identify and characterize the molecular machines of life—the multiprotein complexes that execute cellular functions and govern cell form

**Goal 2:** Characterize gene regulatory networks

**Goal 3:** Characterize the functional repertoire of complex microbial communities in their natural environments at the molecular level

**Goal 4:** Develop the computational methods and capabilities to advance understanding of complex biological systems and predict their behavior

## What Imaging Needs To Provide

Current imaging techniques display a wealth of information about eukaryotic (e.g., human) biological systems over a wide range of length and time scales. Imaging of these systems has led to significant advances in understanding cell function and complex cellular systems (see Fig. ES.1). Microbial systems with their smaller cells, however, present different challenges. New techniques are needed to connect genomic information with microbial functions spatially and temporally in model systems and in their natural environments. These new techniques will drive further advances in all fields of biology.

## Imaging and the Molecular Machines of Life (GTL Goal 1)

Imaging will contribute directly to identifying and characterizing the molecular machines of life, giving a deeper understanding of their relationships. Imaging will help define interactions among proteins and other cellular components in the complex interacting networks of living cells. A real-time, molecular-scale description of protein interactions will reveal metabolic relationships that can be engineered to accomplish DOE missions. High-throughput methods

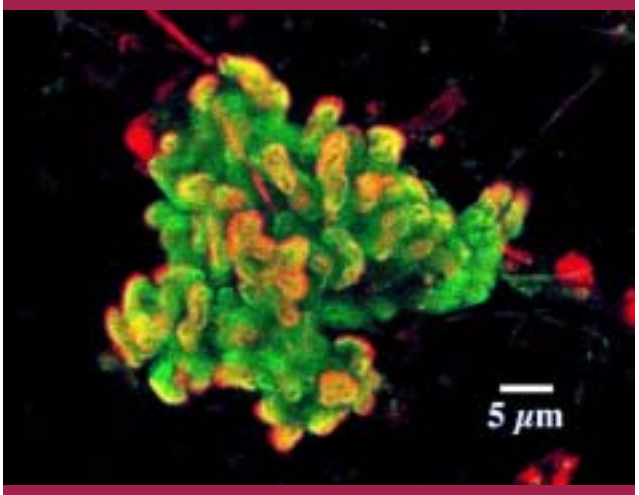



**Fig. ES.1. Confocal laser micrograph of a bacterial microcolony in a river biofilm community. The colony is stained with the electrical-potential–sensitive fluorescent stain JCI; orange regions are areas of high potential.** [Source: John R. Lawrence, National Water Research Institute, Canada]

(e.g., mass spectrometry) for characterizing protein complexes require validation of the existence and function of these complexes in living cells. Imaging can provide that validation.

New imaging methods will be required to define the state of a biological system in response to differing environmental conditions and enable the functional interpretation of traditional analyses of protein complexes. Imaging provides a direct link among the genomes of microorganisms and the atomic structures of the molecular machines that define their functions. Direct observations of protein complexes that comprise these machines will, in turn, provide an important link between genomic information and living cell function. Realizing these potentials will require innovative probes to visualize the structure and dynamics of molecular machines and to locate specific proteins. Substantial innovation will be needed to develop spectroscopies that enable measurements of dynamics (function); microscopies with sufficient spatial resolution and sensitivity to image individual proteins; methods to resolve their atomic structures; and computational methods to acquire, store, access, visualize, and interpret results (see GTL Goal 4).

## Imaging to Characterize Gene Regulatory Networks (GTL Goal 2)

Imaging the location of regulatory proteins in vivo to identify their binding sites in DNA or other cellular structures is needed to understand the primary step of complex gene regulatory networks. The identities of most of these regulatory proteins are unknown. Methods must be developed to interrogate DNA-protein bound pairs on very rapid time scales during the cycles of cell growth and function.

To understand the functions of gene regulatory networks, the dynamic timing of gene expression needs to be known as a function of cell cycle and stimulating signals. This requires development of small fluorescent expression tags that can be imaged without delay in vivo. In particular, we must know the temporal sequence for expression and intracellular distribution of the regulatory proteins themselves to design computational models of the networks.

Knowledge of gene regulatory networks for both microorganisms and eukaryotic systems and imaging of the expression schedule and subsequent distribution of each gene product can provide a basis for understanding the molecular machines of life and their coordinated function in complex microbial communities in natural environments. The gene regulatory networks act as a digital (molecule-by-molecule) computer to specify the identity and level of expression of target genes. Computer models must be developed to enable broad interpretation of experimental results, leading to predictions of biological function (see GTL Goal 4).

## Imaging to Characterize Complex Microbial Communities in Model and Natural Environments at the Molecular Level (GTL Goal 3)

The past decade's advances in techniques for imaging living biological material have revealed that microbial communities are dynamic structured assemblages with compartmentalized (e.g., metabolic) activities. Imaging methods with multiple modalities enable interrogation of these spatially and temporally organized features in a multiplexed manner. Then, to understand life in naturally occurring communities, we must know the identities of the constituent species, their functions in the community, how they perform these functions, and how the communities change over space and time. Attaining these objectives will require development of an advanced suite of probes, imaging devices, and computational methods (see Goal 4). Understanding the complexity of natural systems will require direct study coupled with research on better-defined model systems (see Fig. ES.2).

## Imaging and Developing Computational Methods and Capabilities (GTL Goal 4)

Once images are acquired, new data-processing methods will be required to access and manage image data, enable visualization, and make possible the quantitative analysis of biological systems and their components. Such processing will enable development of predictive computer models that will be essential for addressing DOE missions.

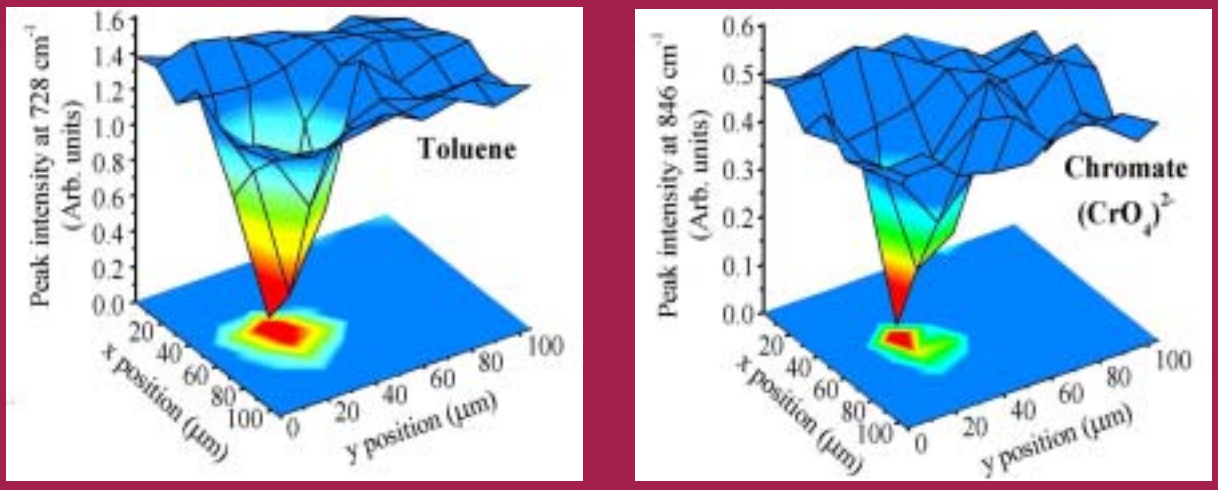



**Fig. ES.2. Synchrotron infrared images showing a small colony of natural bacteria. A common organic contaminant, toluene, was used to accelerate the reduction of a carcinogenic form of chromium (chromate $CrO_4^{-2}$) to its environmentally safe form.** [Source: H.-Y. Holman et al., "Real-Time Characterization of Biogeochemical Reduction of Cr (VI) on Basalt Surface by SR-FTIR Imaging," *Geomicrobiology Journal* **16**(4), 307–24 (©1999). Reproduced by permission of Taylor & Francis, Inc. (www.routledge-ny.com)]

# Recommended Investment Strategy and Timeline for Accomplishments

We recommend a GTL imaging program that integrates technical approaches and biological needs. This program should draw on existing capabilities from other disciplines and advance existing methods to address the unique requirements of the GTL program. It should immediately initiate research toward the most significant imaging challenges, including innovative research with high risks and high potential payoffs. The GTL imaging program should include single- and multi-investigator projects and multi-institutional research programs, funds to develop and maintain essential capabilities at DOE national laboratories, and investments in education. Bringing together scientists from different backgrounds who will create interdisciplinary approaches to important technical challenges will be critical.

## Expected Time Frame for Accomplishments

- Within 2 to 3 years: Capture, integrate, extend, and apply existing imaging technologies to microbes and microbial communities.
- Within 5 years: Identify fundamental bottlenecks and limitations on reaching the potential of existing imaging technologies; develop capabilities for initial conversion of these potentials to practice; develop a pathway for addressing the most challenging potentials.
- Within 20 years: Develop technologies to quantify functions of machines, cells, and communities in real time and in situ with minimal perturbation to the systems; apply these technologies to relevant organisms and establish predictive models.

## Impact of Imaging on GTL Goals and DOE Missions

Coupled with computational modeling and the use of genomic, proteomic, and related analytical information, imaging will quantify functions of machines, cells, and communities, which in turn will enable the use of microbial systems to solve problems in DOE mission areas. Visualization and quantitative image analysis of biological systems and their components provide a level of understanding of complex systems that cannot be obtained in any other manner.